Introduction

For this assignment, I analyzed subjects' MRI scan data across two visits. The dataset also included details on subjects' age, gender, socioeconomic status, and education level. The goal was to examine how the scan parameters changed over time, considering factors like group and gender, and aimed to investigate how within-subject factors influenced changes in MRI scan parameter values, while also considering between-subject factors.

Data Pre-Processing

First, I treated the dataset for any missing values, replacing them with their mean values from their respective columns. To access the variables more efficiently, I renamed the relevant column names and labels of the values in the 'Gender' column for better readability.

Exploratory Data Analysis

To study the overall distribution of the values of the information from the subjects' MRI scans, I constructed a summary statistics output based on the following columns - 'Age', Mini Mental State Examination ('MMSE'), Clinical Dementia Rating ('CDR'), Estimated Total Intracranial Volume ('eTIV'), Normalize Whole Brain Volume ('nWBV'), Atlas Scaling Factor ('ASF'). The frequency tables showed the highest count for subjects under the 'Nondemented' category followed by 'Demented' and 'Converted', with female 'nondemented' subjects having the highest count, and an overall higher count of female subjects than male. Following this, I created a series of visualizations of the distributions of the above variables, disaggregated by gender/group categories using density plots and boxplots, resulting in varying inferences across different parameters based on their categories. The data showed subjects from the 'Converted' group were of slightly higher ages (80-90) compared to the other two groups, with a majority of subjects falling under the three groups of ages between 60-90, with an overall mean of 76 years of age.

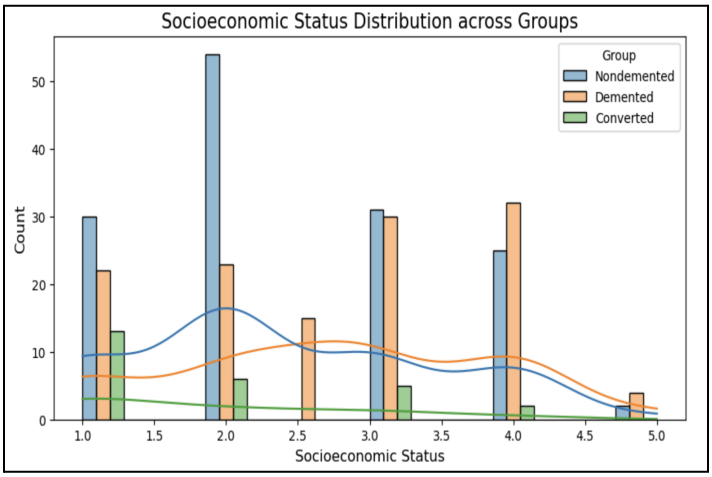


Figure 1. Socioeconomic Status Distribution across Groups

The socioeconomic status of most of the subjects lay in a similar range across the group categories, with the exception of 'nondemented' subjects, who had the highest count, with a status level of 2, as seen in Figure 1. Further, other graphs of the data also showed female subjects having a higher education level than male subjects. Lastly, in a series of comparative grouped graphs of the scan parameter values across the two visits (Visit 1 and Visit 2), the majority of the data showed a common trend for lower values of the parameters in Visit 2 than in Visit 1.

For reference:

Hypotheses of Statistical Tests used for testing the assumptions -

- **Mauchly's test of sphericity (For testing sphericity)**
  - **Null Hypothesis -** Variances of differences of independent variable values across both visits are equal
  - **Alternative Hypothesis -** Variances of differences of independent variable values across both visits are not equal
- **Shapiro Wilk Test (For testing normality)**
  - **Null Hypothesis -** Data is drawn from normal distribution
  - **Alternative Hypothesis -** Data is not drawn from normal distribution
- **Levene's Test (For testing homogeneity of variances)**
  - **Null Hypothesis -** Samples from populations have equal variances
  - **Alternative Hypothesis -** Samples from populations do not have equal variances

Two-Way Mixed ANOVA - I

RQ: Does the group categorization have any impact on nWBV (Normalize Whole Brain Volume) across the two visits?

Null hypothesis: Means of nWBV values among different groups across the visits are equal

Alternative hypothesis: There is at least one difference in the means of nWBV values among different groups across the visits

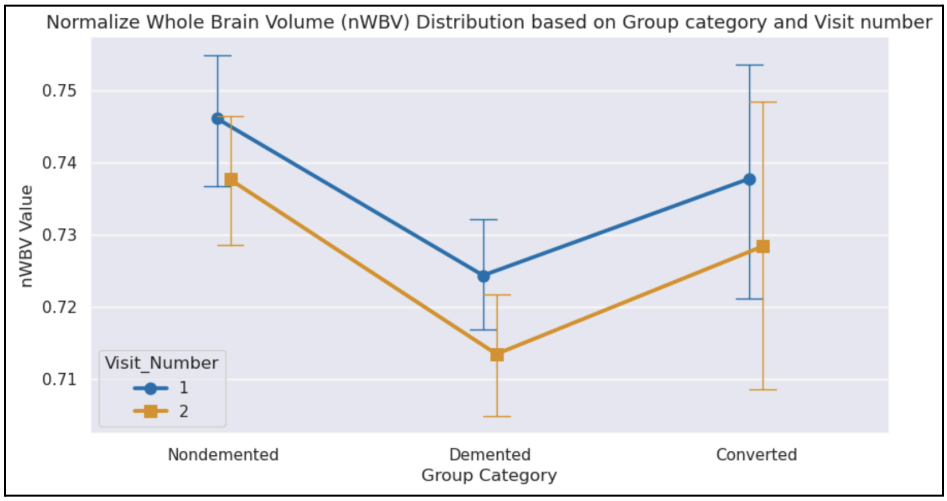


Figure 2. Plot for Mixed ANOVA Model - I

The plot clearly shows a lower value for nWBV values for Visit 2, compared to Visit 1 across all three groups, as seen in Figure 2.

Results from ANOVA:

Main Effect 'Group' - p-value = 0.002, F-statistic = 6.712

Main Effect 'Visit Number'- p-value < 0.001, F-statistic = 94.251

Interaction Effect - p-value = 0.219, F-statistic = 1.534

Since the p-values < 0.05, we reject the null hypothesis for the main effects and conclude that the group categories and visit numbers do have a significant effect on the subject's nWBV values.

The results also show the p-value for the interaction effect as > 0.05, therefore, we do not reject the null hypothesis for the interaction effect of group category and visit numbers, and conclude visit numbers do not impact the nWBV values based on the subject's group.

Testing Assumptions

- *Assumption 1: Assumption of sphericity* - Mauchly's test of sphericity concludes that the data meets the assumption of sphericity since the p-value > 0.05, and therefore, we do not reject the null hypothesis.
- *Assumption 2: Test for Normality (Each level of within-subjects factor is normally distributed)* - The Shapiro-Wilk Test concludes that the data is normally distributed since the p-value > 0.05 at both levels of visits, and therefore, we do not reject the null hypothesis.
- *Assumption 3: Homogeneity of Variances* - Using Levene's Test at each level of visit numbers resulted in both p-values > 0.05. Therefore, we do not reject the null hypothesis and conclude that the variances are homogeneous at both levels of the visit.

Two-Way Mixed ANOVA - II

RQ: Does the gender categorization have any impact on eTIV (Estimated Total Intracranial Volume) across the two visits?

Null hypothesis: Means of eTIV values among different genders across the visits are equal

Alternative hypothesis: There is at least one difference in the means of eTIV values among different genders across the visits

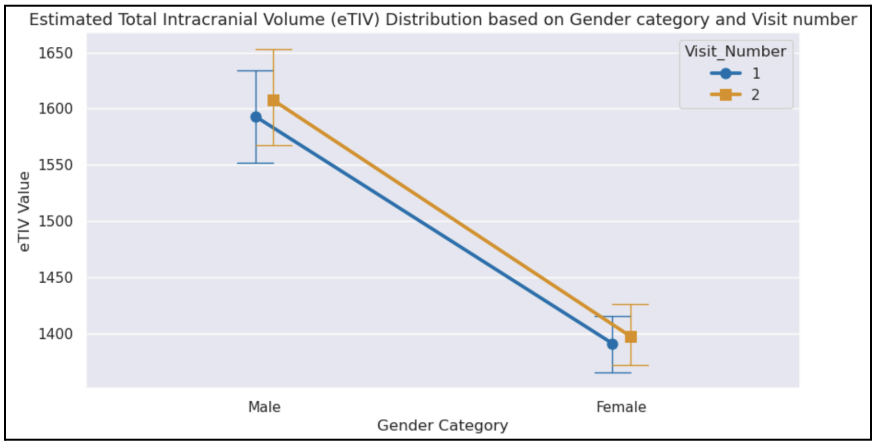


Figure 3. Plot for Mixed ANOVA Model - II

The data shows a small increase in eTIV values in Visit 2, compared to Visit 1 for both genders, as seen in Figure 3.

Results from ANOVA:

Main Effect 'Gender' - p-value < 0.001, F-statistic = 72.906

Main Effect 'Visit Number'- p-value = 0.003, F-statistic = 9.208

Interaction Effect - p-value = 0.531, F-statistic = 0.394

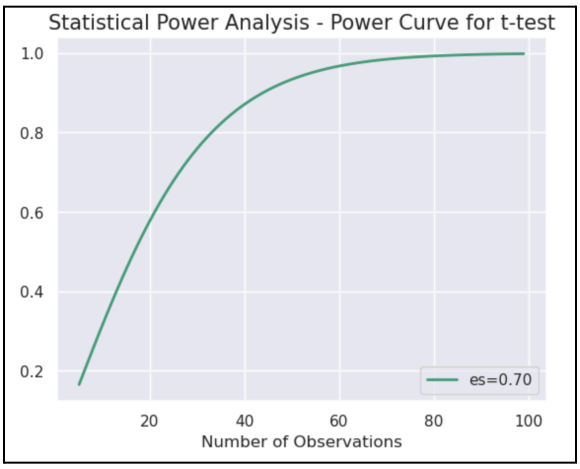
Since the p-values < 0.05, we reject the null hypothesis for the main effects and conclude that the gender categories and visit numbers do have a significant effect on the subject's eTIV values.

The results also show the p-value for the interaction effect as > 0.05, therefore, we do not reject the null hypothesis for the interaction effect of gender and visit number, and conclude visit numbers do not impact the eTIV values based on the subject's gender.

Testing Assumptions

- *Assumption 1: Assumption of sphericity* - Mauchly's test of sphericity concludes that the data meets the assumption of sphericity since the p-value > 0.05, and therefore, we do not reject the null hypothesis.
- *Assumption 2: Test for Normality (Each level of within-subjects factor is normally distributed)* - The Shapiro-Wilk Test concludes that the data is not normally distributed since the p-value < 0.05 at both levels of visits, and therefore, we reject the null hypothesis.
- *Assumption 3: Homogeneity of Variances* - Using Levene's Test at each level of visit numbers resulted in both p-values < 0.05. Therefore, we reject the null hypothesis and conclude that the variances are not homogeneous at both levels of the visits.

Statistical Power Analysis



Based on the pre-defined parameters for the power analysis of power = 0.91, alpha = 0.05, and effect size = 0.7, I calculated the estimated sample size using t-tests, applicable to both RQs. The results from the code suggested a minimum sample size of 45 for each group to result in having a significant p-value in the t-test. This is similarly observed in the power curve using the same above parameters and an array of sample sizes ranging from 5 to 100, as seen in Figure 4.

Figure 4. Statistical Power Curve Plot

Conclusion

The results suggest significant changes in the scan parameter values between Visit 1 and Visit 2, with differing trends across group and gender categories. Although both the mixed ANOVA models did not reveal a significant interaction effect, the main effects had a significant impact on the parameter values. It would be interesting to expand this study by examining other within-subject factors besides visit numbers to see how they affect scan values and to identify any consistent patterns in the data.

References -

- Python Graph Gallery - https://python-graph-gallery.com
- Stack Overflow - https://stackoverflow.com
- Geeks for Geeks - https://www.geeksforgeeks.org/

- Seaborn Documentation - https://seaborn.pydata.org/archive/0.11/tutorial/categorical.html
- Pandas DF Documentation - https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.html
- Mixed ANOVA using Python and R (with examples), https://www.reneshbedre.com/blog/mixed-anova.html
- Repeated measures ANOVA in Python, https://raphaelvallat.com/pingouin.html
- Seltman, H. J. (2018). Experimental Design and Analysis, Department of Statistics at Carnegie Mellon